A Critical Study of Artificial Intelligence in Healthcare: Prospects and Perils

Thalia Bueno a*
 Department of Social Science, York University, Toronto

ABSTRACT: The modern era has ushered the proliferation of new technologies, especially witnessed in the emergence of the nascent artificial intelligence (AI) sector. The use of AI is largely multifaceted, proving useful in various industries such as healthcare - however, it may also allow for deleterious effects to occur. The use of AI in healthcare settings can work to extend and augment the quality of patients’ lives. Notwithstanding this, health AI enshrines various perils including the lack of patient privacy, and algorithm bias - particularly on marginalized and racialized communities. This is ultimately compounded by the absence of ethical frameworks governing the usage of AI in healthcare settings. Specifically, this article seeks to explore whether or not the use of health AI is a potential prospect or peril; considering its duality. To investigate the nuances of health AI, this article will utilize an interdisciplinary approach – drawing upon research from domains such as: sociology, socio-legal and socio-medical climates. This study finds that health AI remains a greater prospect – as it reinforces the quality and elongates the duration of the human lifespan. It concludes with a call to action to inform the success of health AI in praxis: namely, the need to incorporate the aforementioned topics within medical pedagogy and ethical frameworks.

KEYWORDS: sociology of AI, social determinants of health, bioethics, artificial intelligence, medical education
A Critical Study of AI in Healthcare (Bueno, Thalia)

Introduction

As the world continues to evolve and morph, society faces many prospects and perils that can safeguard or threaten the mere existence of humanity. Imminent existential threats are derived from anthropogenic risks, which are, “induced entirely or predominantly by human activities and choices”.¹ In recent years, the nascent AI sector has begun to grow, which has led society to consider if it will a prospect or peril to the future of humanity. Ord writes that the purpose of AI is to, “build machines rival[ing] humans in their intelligence”.² It is preeminent to acknowledge that the use of AI should not inherently be labelled as ‘good’ nor ‘bad’. Rather, the categorization of AI into this dichotomy is dependent on implementation, oversight, and mitigating measures in praxis. AI can be utilized for a multitude of purposes, and has the potential to augment healthcare experiences.

When utilized in healthcare, AI can have global, wide-spread implications as it can positively contribute to the overall life expectancy and quality of life experienced by any given nation’s citizens. However, it can also be weaponized and contribute to inequitable healthcare treatment of marginalized communities, drastically decreasing life expectancy and quality of life. The relevance of AI in healthcare and its usage by medical practitioners are inherently collateral objectives. It is ultimately supported by medical professionals’ sheer obsession on the conquest of death, which seeks to avoid death for as long as humanly possible utilizing modern technology.³ In the healthcare domain, the use of AI can pose as an existential risk due to the lack of patient privatization in data collection, presence of algorithmic bias, and absence of ethical/legally binding framework. This type of unaligned AI has the potential to be weaponized against nations, and impact aspects of globalization. Alternatively, AI in healthcare can also be a prospect if it has a security factor such as laws and regulations that protect patient privacy in data collection, as it can result in sustainable innovation by removing bias. This paper will explore if AI in healthcare should be considered a greater prospect or peril due to the nature of its duality.
Methodology

To study this, this paper will utilize a comparison method to provide a robust explanation as to how AI can present itself as a prospect and/or peril. The objective of this paper is neither a systematic review of the literature nor a scoping review. Furthermore, it will utilize an interdisciplinary lens that draws from different domains such as, sociology (social determinants of health), socio-legal, and socio-medical climates. Quantitative and qualitative research will be explored in order to gather both numerical and non-numerical data which can be used to provide further explanation. Evidence will be mainly sourced from peer-reviewed journal articles, institutional reports, books and contemporary alternative sources. Hypothetical and empirical case studies will be utilized to conceptualize and demonstrate the impact of AI in healthcare.

Advances Towards Ethical Practices of AI

Notwithstanding the nascency of AI, many states have become increasingly aware of the vast potential that AI holds and are now in the beginning stages of policy development. Canada’s Bill C-27, *An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act* and to make consequential and related amendments to other Acts, serves as a prime example. Additionally, others have already passed AI-related policies. On March 13, 2024, the European Parliament (EP) passed the *Artificial Intelligence Act*, which provides a human-rights focused legal framework to regulate the market and augment public trust in AI. Advancement in proposed Canadian and EP policies are beneficial as they protect and mitigate the potential ill effects that AI may have on its citizens. However, it is equally critical to note that the AI sector is subject to rapid advancement, and thus, may require more a complex and specific legal framework over time.

Similarly, advances in AI scholarship have fostered progress within medical school curricula. As part of their educational
programs, at least 11 medical schools in Canada, the United States, and South Korea offer AI related programming. Such progress is commendable as it serves to introduce medical students to health AI in praxis by providing instruction on navigating rudimentary technological advancements in modern medicine. It remains critical that such programs are disseminated to additional medical schools to provide all students with robust and equitable educational opportunities.

The Intersections of Privacy and AI Predictions

AI data collection in healthcare poses a duality as it can provide for more precise predictions, yet have implications in privacy breaches. A South Korean research company, Seegene, used AI to complete a big data analysis using RNA information found in online datasets to develop polymerase chain reaction (PCR) test kits. RNA datasets used by AI can serve as a place for scientists to collect large amounts of information. This allowed for researchers to identify predictions in RNA activity, which enabled the development of PCR tests. These tests have played a huge part in greater society as they were used for public health measures during the COVID-19 pandemic. Typically, AI’s governance of privacy and use of human data collection varies by bureaucratic structure (often by nation, region, and/or municipality). This means that human data may or may not be permissible to utilize for health AI development and research. Bak et al. notes that data that is used in health AI cannot guarantee full anonymization due to the inclusion of genetic sequences. Data collection for datasets oftentimes run on the premise that a patients’ identifiers are protected. However, this is not inherently the case as full anonymization of data is not possible due to the presence of genetic identifiers. The conundrum of privacy in data collection emulates the poor and lack of legally-binding regulations, which can pose as an overall e-risk.

Ord touches on how technological advancement can increase societal potential with regards to the duration of the human lifespan. Through utilizing AI in healthcare, computers are able to analyze
datasets to efficiently produce far superior results than researchers alone. The importance of technological advancement in healthcare is especially imperative as medical professionals are on the conquest of death. The application of AI extending the duration of human life is exemplified in the aforementioned example of the COVID-19 pandemic. It allowed for healthcare professionals to effectively and efficiently diagnose and treat cases of COVID-19, in turn, extending the duration of ill patients. Moreover, the use of AI in healthcare can augment societal potential by preserving lives. Annually, the use of Health AI has the potential to save an upwards of 403,000 lives. Denoting these statistics, it is assumed that patients will have augmented treatment outcomes and quality of life. AI has the ability to greatly impact and extend human life, however the implications surrounding a lack of privacy can inhibit its advancement.

Notwithstanding these prospects, the lack of privacy regarding AI in healthcare settings can potentially enable more privacy breaches through data leaks. Alder indicates that the lack of reporting transparency makes it difficult to quantify the magnitude of healthcare data breaches, and concur that the use of AI will lead to increased number of data breaches. The absence of legal framework and regulatory body, results in a lack of statistics on the frequency of current and projected data breaches. Datasets can serve as a major identifier for individuals due to genetic markers, compromising one’s identity. While the importance of privacy is not lost, the integration of AI both works to save lives, and specifically, the use of RNA datasets allowed for accelerated research as results were compiled more efficiently. This ultimately enabled society to fight against a global health crisis, the COVID-19 pandemic. It is important to note that concerns surrounding privacy can be mitigated via a security factor, such as the implementation of control regulations to ensure that access is restricted.

The Crossroads of Algorithms and Marginalized Communities

The use of algorithms in health AI can provide for superior reasoning which can augment clinical decisions. However, it can also
provide for algorithm bias to flourish as many health AIs fall prey to utilizing small input data group(s). The use of AI can remove potential medical and/or social biases a physician may have, thereby leading to improved patient prognosis. A case study illustrates how AI can be used in clustering techniques to determine categories of patients with sinus rhythm and atrial fibrillation responding to β-blockers. In this vignette, AI can be regarded as superior in human ability as it efficiently organizes and clusters patients on their relative response rate to β-blocker therapy compared to manual input. As a result, physicians are able to mitigate any potential preconceived notions of health of the patient, which in effect remove social and medical biases. This can consequently improve the lives of racialized, gendered and classified bodies, in a systematically-affirming manner. These simple tasks completed by AI enables physicians to engage in more direct aspects of patient care as it alleviates occupied time doing cumbersome tasks.

However, it is integral to note that algorithms are not inherently neutral; social and medical bias can also manifest in AI through forms of algorithmic bias. Due to the presence of algorithmic bias, the effectiveness of such AI is reduced. Machine learning can be defined as a segment of AI which utilizes algorithms to imitate the human learning experience, and honing accuracy. One study found that an x-ray AI technology only drew aspects of machine learning from racially and ethnically monolithic datasets. It is important to realize that AI is not inherently neutral as popular belief stipulates, and that rather often includes instances of systemic discrimination. Traditionally, marginalized and objectified communities in healthcare such as bodies of colour and young women end up underdiagnosed. This could consequently work to perpetuate this marginalization, and further contribute to health inequity.

In clinical application, the presence of algorithmic bias, the effectiveness of the AI becomes self-inhibiting as it can dramatically worsen patient outcomes. Kamulegeya et al. determined that an AI imaging app developed by First-Derm had a 17% diagnostic accuracy on Fitzpatrick V and VI skin types versus a 69.9% accuracy rate on type I and II skin, resulting in an underdiagnosis of fungal infections
in the former group.\textsuperscript{16} In such scenario, marginalized and specifically racialized communities did not receive an adequate diagnosis utilizing AI; allowing their skin infections to persist. By utilizing monolithic datasets, AI continues to perpetuate this marginalization which results in health inequity – by worsening clinical outcomes in racialized communities. To combat the issue of algorithmic bias, a security factor must be used, such as incorporating regulation that requires the use of diverse datasets in machine learning for healthcare AI.

In chapter eight, Ord discusses how society has the potential to extend the duration of human lives through technology development.\textsuperscript{17} Through further research, using AI in healthcare domains can assist society in achieving a longer lifespan. In Canada, the shortage of family physicians impacts primary care, Li et al. indicates that, "every additional FP per 10,000 people increases life expectancy by 51.5 days."\textsuperscript{18} A virtual primary care clinic, K Health, had diagnostic algorithms that were trained on a dataset of 2 million people, first yielding a 96.6\% accuracy rate, with nominal discrepancies across demographics.\textsuperscript{19} Used correctly, the collection of diverse age, ethnical, and racial datasets allowed for rigorous and robust algorithm training, which engaged in systemically affirming care. By alleviating some of the workload on primary care physicians, this allows them to provide more efficient and effective care. This particular AI could encourage more individuals to access primary care to address their current health concerns, thereby potentially extending the duration of their lives. Simultaneously, it may also serve as an opportunity to treat underserved demographics (i.e. individuals with health conditions, or from rural communities) through providing remote primary care.

**Situating Ethics in Health AI: Research, Implementation and Education**

Lastly, the lack of an ethical framework of health AI can allow for accelerated development, but can also contribute to unsustainable practices and development of health AI. As of 2018, the
FDA started approving health AI devices for human usage – as of 2021, there are over 160 FDA approved devices. This has insinuated the rapid development of health AI as researchers look to gain FDA approval. The nature of which these devices are approved can be alarmingly inappropriate as more medical professionals, specifically physicians will be expected to know how to use them. Katznelson and Gerke note that there is a lack of education on AI ethics for medical school students. As new generations of medical doctors enter a new generation of medicinal practices, they will be at a disadvantage. Health AI requires a robust understanding before it can be implemented in healthcare appropriately, medical school graduates will remain unaware of ethical conundrums at play. This has the potential to create unique and precarious situations, as the lack of ethical knowledge surrounding the usage of AI can inhibit sustainable practices and contribute to the overall e-risk of health AI.

In application, the implications of the absence of ethical framework often manifests within the complex construction and usage of the AI itself. Corti, a machine learning software that analyzes emergency phone calls to determine if the caller has cardiac arrest; however, Corti’s inventor does not fully understand how the algorithm deducts decisions, a phenomenon known as a black box. The usage of this health AI has the potential to save lives by accurately identifying the health status of the caller, and work to expedite emergency services. The absence of ethical frameworks promotes potentially hazardous activities, as witnessed in the fact that healthcare practitioners and inventors do not fully understand the deductive reasoning of the algorithm. Due to this, healthcare professionals remain limited in their interpretive scope to identify the logical reasoning behind the algorithm which may incidentally reinforce notions of bias.

An ethical framework on Health AI could focus on principles such as autonomy and justice. Beauchamp and Chambliss conceive the principle of justice by focussing on distributive justice which is the entitlement to just, equitable and apt treatment, through accounting for unique social identities. Noting that AI is subject to unprecedented growth, the aforementioned principle could work to
centre and acknowledge the potential peril of algorithmic bias (i.e. social, gendered, and racial bias) in the context of treatment. This could work to inform the developers and researchers of health AI, which can mitigate the development of AI on the basis of social stereotypes. Furthermore, ensures patient autonomy could work to strengthen the use of health AI in treatment and research. Autonomy refers to the idea that patients are able to hold their own perception and make decisions based on their values and beliefs. The notion of autonomy could work to strengthen the involvement of patients in determining the integration of advanced health technology in the course of their treatment. It also reinforces the need to attain informed consent when collecting data for pioneering research of health AI. Conceptualizing specific ethical principles (i.e. transparency) warrants further discussion as there are various factors to consider (i.e. cultural context).

Further, Ord discusses how risk landscape of a particular e-risk can be calculated utilizing this formula as seen in (1):

\[
\text{Cost effectiveness} = \text{Importance} \times \text{Tractability} \times \text{Neglectedness}
\]  

The use of this formula has its limitations: 1) e-risks are quantifiable, and 2) the importance of AI is subject to individual interpretation. While it can be argued that the use of a formula remains inappropriate to calculate e-risk potential. Consequently, for the purposes of this paper, the intention of using this formula is not to assign a particular numeral value to the particular risk, but rather provide an overview of the e-risk. Prescribing a more qualitative approach to discussing AI’s risk factors would be most effective. However, if the main concepts from the formula are utilized, it is evident that AI is not adequately addressed and remains a substantial e-risk.

Therefore, in order to address the e-risk posed by AI in healthcare, society must seek active international cooperation. The use of AI has the potential to transcend borders of the state, and can be developed transnationally, therefore, there is a need to focus on global governance. The usage of health AI is not strictly limited to
one state, and has the potential to radically transform patient experiences in healthcare. In 2023, the UN launched the Advisory Body for AI which aims to augment advice for the international governance of AI. The integration of an international advisory body is notable, and with further development has the potential to develop and/or become a governing entity. While global governance could provide for overall oversight, it is important to ensure multi-tier governance to address gaps that may not be accounted for in an international body, and hold actors within the state accountable. Adopting a multi-level governance approach could be beneficial in ensuring regulation is met at regional, national, and international levels. The importance of multi-level governance for health AI is integral as this can potentially provide for sustainable research and development.

**Conclusion**

This paper ultimately worked to explore if AI in healthcare is a greater prospect or peril due to the nature of its duality. As counterclaims discuss AI’s lack of privacy, algorithm bias and absence of ethical framework, these can be easily mitigated. The implementation of AI in healthcare can allot for privacy breaches as information is utilized to create datasets on the premise that the information would be fully anonymized. However, Bak et al. illustrate that the concept of anonymity is not possible. While protecting patient privacy remains of utmost concern, these can be mitigated using security factors surrounding who is privileged in accessing and isolating genetic identifiers. These datasets can provide a wide breadth of genetic information which can be utilized in AI, and ultimately work to drastically extend a patient’s life. While AI datasets can be racially and ethnically exclusive, leading to an underdiagnosis in marginalized patients, this can be offset by collecting more diverse data. The collection of diverse data can be ensured using the security factor of drawing upon ethical regulations that include notation on the requirements to compile and utilize diverse datasets in health AI.
The potential of extending humanity’s lifespan, and ameliorating the quality of life, in perspective, makes the peril appear negligible. These consequences can also be addressed via a security factor - the implementation of security factors such as international cooperation in implementing ethical, legally-binding frameworks. This will work to ensure that the scope of the health AI is within set parameters and provides equitable, fair, and just healthcare for all. Through the robust analysis in earlier discussion, it is evident that AI remains a far greater prospect than peril as it has enduring effects such as improving the duration and quality of human life. This is established as it can provide better predictions through machine learning, mitigate physician bias, and can be grounded by the implementation of ethical frameworks. Overall, it is imperative to engage in critical study to implement legally-binding ethical framework and regulations on an international level to mitigate potential e-risks that AI in healthcare poses.
Notes


3 Dagmara Woronko. “Technology and Medicalization” (PowerPoint Slide, York University, 2022)


9 Ord. The Precipice: Existential Risk, 217-241

10 Dagmara Woronko. “Technology and Medicalization”


A Critical Study of AI in Healthcare (Bueno, Thalia)


Bibliography


A Critical Study of AI in Healthcare (Bueno, Thalia)


- 18 -


Woronko, Dagmara. Technology and Medicalization.” PowerPoint Slide, York University, 2022

A Critical Study of AI in Healthcare (Bueno, Thalia)